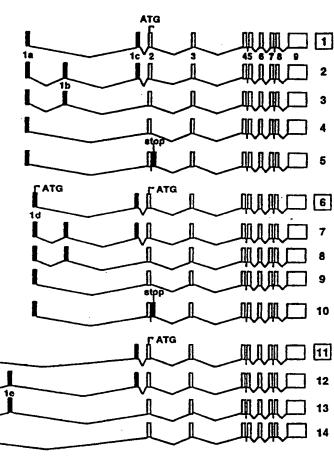


B.



C.

Transcript 1: MEAMA ASTSL PDPGD FDRNV PRI DBD 427aa

Transcript 6: MEWRN KKRSD WLSMV LRTAG VEEAF GSEVS VRPIR RAPLG STYLP PAPSG MEAMA ASTSL PDPGD FDRNV PRI DBD 477aa

Transcript 9: MEW RNKKR SDWLS MVLRT AGVEG MEAMA ASTSL PDPGD FDRNV PRI DBD 450aa

FIGURE 1

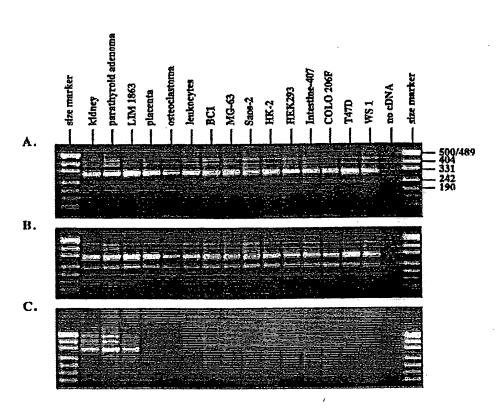


FIGURE 2

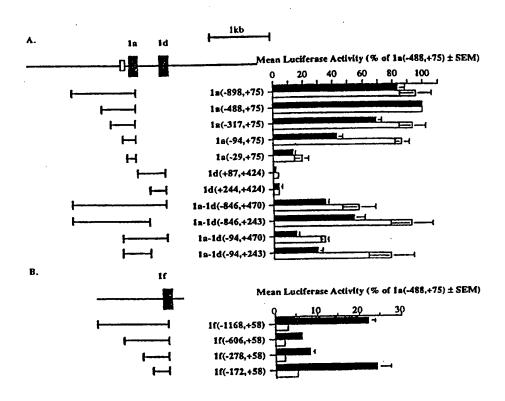


FIGURE 3

- C. 5'...tgtttttagAGGCAGCATGAAACAGTGGGATGTGCAGAGAGAAGATCTGGGTCCAGTAGCTCTGACACTCCTCAGCTGTAGAAACCTTGACAACTCTGCACATCAGTTGTACAATGGAACGGTATTTTTTACTCTTCATGTCTGAAAAGGCTATGATAAAAGATCAAgtaagatatt...3'

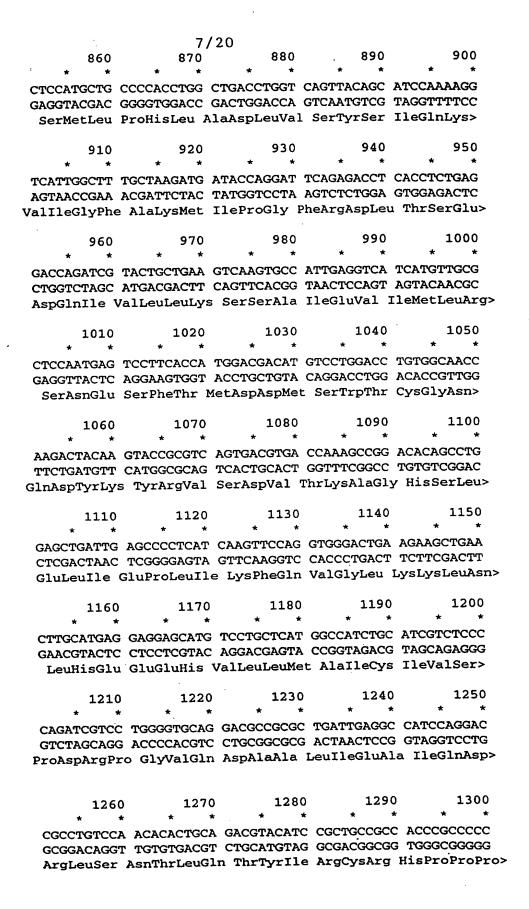
FIGURE 5 TRANSCRIPT 6

(Sequence Range: 1 to 1463)

ange: 1 to 1	1463)			
4.0	20	30	40	50
10	20	* *	. * *	* *
	==	CGCCTTGGCA	тесьстесьс	GAATAAGAAA
GITTCCTTCT	ACACACCCCC	GCGGAACCGT	ACCTCACCTC	CTTATTCTTT
CAAAGGAAGA	AGACAGCCCC	GCGGAACCG1	MetGluTrpAr	g AsnLysLys>
		•		9
60	70	80	90	100
* *	* *	* *	* *	* *
AGGAGCGATT	GGCTGTCGAT	GGTGCTCAGA	ACTGCTGGAG	TGGAGGAAGC
TCCTCGCTAA	CCGACAGCTA	CCACGAGTCT	TGACGACCTC	ACCTCCTTCG
ArgSerAsp '	TrpLeuSerMe	t ValLeuArg	ThrAlaGly	ValGluGluAla>
	_		•	
110	120	130	140	150
* *	. * *	* *	* *	* *
CTTTGGGTCT	GAAGTGTCTG	TGAGACCTCA	CAGAAGAGCA	CCCCTGGGCT
GAAACCCAGA	CTTCACAGAC	ACTCTGGAGT	GTCTTCTCGT	GGGGACCCGA
PheGlySer	GluValSer	ValArgProHi:	s ArgArgAla	ProLeuGly>
			400	200
160	170	180	190	200
* *	* *	* *		
CCACTTACCT	GCCCCCTGCT	CCTTCAGGGA	* COMCOCAMI	CCCCCCCTCC
GGTGAATGGA	CGGGGGACGA	GGAAGTCCCT	ACCICCGIIA	t AlablaCar>
SerThrTyrLe	u ProProAla	Prosercty	Mergrunrame	t AlaAlaSer>
210	220	230	240	250
210			* *	
			CGGAACGTGC	CCCGGATCTG
TCA ACCCACC	CIGACCCICC CACTGGGACC	TCTGAAACTG	GCCTTGCACC	GGGCCTAGAC
Threatle	ProAspProGl	v AspPheAsp	ArgAsnVal	ProArgIleCys>
IIII Ser nea	rionspire-	.,	_	
260	270	280	290	300
* *	. * *			k * *
TGGGGTGTG	GGAGACCGAC	CCACTGGCTT	TCACTTCAA	r GCTATGACCT
ACCCCACACAC	CCTCTCTCCCTY	CGTGACCGAA	AGTGAAGTT	A CGATACTGGA
GlyValCys	s GlyAspArg	AlaThrGlyPh	e HisPheAs	n AlaMetThr>
310	329	0 330		_
*	* *		* *	
GTGAAGGCT	G CAAAGGCTT	C TTCAGGCGA	A GCATGAAGC	G GAAGGCACTA
CACTTCCGA	C GTTTCCGAA	G AAGTCCGCT".	COTACTICG	C CTTCCGTGAT
CysGluGlyC	ys LysGlyPh	e Pheargarg	Sermethysa	rg LysAlaLeu>
		ი 380	n 39	0 400
36	0 37 * *		* *	* * *
*	~ ~ ~ *			G ACAACCGACG
TTCACCTGC	C CCTTCAACG	C CCTGACGGC	G TAGTGGTTC	C TGTTGGCTGC
AAGTGGACG	DroPheAcnC	lv AspCysAr	g IleThrLys	AspAsnArgArg

PheThrCys ProPheAsnGly AspCysArg IleThrLys AspAsnArgArg>

410	420	430	440	450
* *	* *	* *	* *	* *
CCACTGCCAG	GCCTGCCGGC	TCAAACGCTG	TGTGGACATC	GGCATGATGA
GGTGACGGTC	CGGACGGCCG	AGTTTGCGAC	ACACCTGTAG	CCGTACTACT
HisCvsGln	AlaCysArg 1	LeuLysArgCys	ValAspIle	GlyMetMet>
			_	
460	470	480	490	500
* *	* *	* *	* *	* *
AGGAGTTCAT	TCTGACAGAT	GAGGAAGTGC	AGAGGAAGCG	GGAGATGATC
		CTCCTTCACG		
LvsGluPheIl	e LeuThrAsp	GluGluVal (GlnArgLysArg	g GluMetIle>
510	520	530	540	550
* *	* *	* *	* *	* *
CTGAAGCGGA	AGGAGGAGGA	GGCCTTGAAG	GACAGTCTGC	GGCCCAAGCT
CACTTCCCCT	TCCTCCTCCT	CCGGAACTTC	CTGTCAGACG	CCGGGTTCGA
Leulvehra	LveGluGluGl	n AlaLeuLvs	AspSerLeu	ArgProLysLeu>
Dearlyshig	nyadiudiudi	a manean,		
560	570	580	590	600
* *	* *	* *	* *	* *
CTCTCACCAC	CAGCAGCGCA	TCATTGCCAT	ACTGCTGGAC	GCCCACCATA
CACACTCCTC	CTCCTCCCT	AGTAACGGTA	TGACGACCTG	CGGGTGGTAT
CorCluGlu	Gleclececi	IleIleAlaIl	e LeuLeuAsp	AlaHisHis>
261.010010	GillGilmig	IICIICIICII	C 200200	
610	620	630	640	650
* *	* * *	* *	* *	* *
A C A C C TTA C C A		ጥርርርልርጥጥርጥ	GCCAGTTCCG	GCCTCCAGTT
MCMCC I ACCA	CCCCACCIAC	ACCCTGAAGA	CGGTCAAGGC	CGGAGGTCAA
T-robbedici	n Prombrair	SeraenPhe	CvsGlnPheAr	g ProProVal>
DASIMILATE	b Florming	Delubline	0,00	3
660	670	680	690	700
* *		:	* *	* *
ഗാനാനവും മുന	a Amerimena	CAGCCATCCT	TCCAGGCCCA	ACTCCAGACA
CCACACTTA	TACCACCTCC	CTCGGTAGGA	AGGTCCGGGT	TGAGGTCTGT
Aratal Aan	. IACCACCICO	v SerHisPro	SerArgPro	AsnSerArgHis>
Argvarasii	Apparagrad	y bernistre	Deringin	
710	720	730	740	750
* '1\	* * 120	, . * *		* * *
CACTCCCAG	~ mmcmcmccc	2 አርጥርርጥርርጥር	CTCCTGCTC	GATCACTGTA
CHC1CCCAG	Z NACACACCC	TCACCACCAC	GAGGACGAGI	CTAGTGACAT
When Deep Co.	AAGAGACCC) rongonound	r SerCvsSei	AspHisCys>
Thrprose	r Pheseigly	Aspserserbe	ir berejooe	
76	0 77	780	790	800
*	* *	* * 1	* * 1	k * *
		 ~ ~ X ~ T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T ~ C T	CCTTCTCA	A TCTGGATCTG
TCACCTCTT	C MGACAIGAI	C CACICGICO	r CGAAGAGGT	r AGACCTAGAC
AADADDTDA	o TCTGTACTA	t Acnserser	SerPheSerA	sn LeuAspLeu>
TTelurgers	er wabwerwe	c wabaerner		
81	0 82	ი 830	84	0 850
* 01	* *	•		* * *
		ል	G ACCCTAGAG	C TGTCCCAGCT
ACT CAMOUNA	C WYYCHOWIG	T GGGAAGACA	C TGGGATCTC	G ACAGGGTCGA
TCACTTCTT	CIMMOTOTAL	1 GGGHAGAGA	1 00	LeuSerGlnLeu
CAYCLUCI	Acneerhenn	en Proserva	1 Turrengiu	negzereriineg.



8/20 1330 1340 1350 1310 1320 1330 ** ** ** ** ** ** ** ** ** **
GlySeinia * *
1360 1370 * * * * * * * * * * * * * * * * * * *
ArgSerLeuAsn GluGluHis Selsi 1440 1440 * *
1410 * * * GTGCTCGAAG TGTTTGGCAA * * * * * * * * * * * * * * * * * *

CCTGAGTGCA GCATGAAGCT AACGCCCCTT GTGCTCGAAG TGTTTGGCAA CCTGAGTGCA GCATGAAGCT AACGCCCCTT GTGCTCGAAG TGTTTGGCAA GGACTCACGT CGTACTTCGA TTGCGGGGAA CACGAGCTTC ACAAACCGTT ProGluCys SerMetLysLeu ThrProLeu ValleuGlu ValPheGlyAsn>

1460 TGAGATCTCC TGA ACTCTAGAGG ACT GluIleSer ***> FIGURE 6

9/20

TRANSCRIPT 9 (Sequence Range: 1 to 1382) 50 40 30 20 MetGluTrpArg AsnLysLys> 100 90 80 AGGAGCGATT GGCTGTCGAT GGTGCTCAGA ACTGCTGGAG TGGAGGGGAT 70 TCCTCGCTAA CCGACAGCTA CCACGAGTCT TGACGACCTC ACCTCCCCTA ArgSerAsp TrpLeuSerMet ValLeuArg ThrAlaGly ValGluGlyMet> 130 GGAGGCAATG GCGGCCAGCA CTTCCCTGCC TGACCCTGGA GACTTTGACC CCTCCGTTAC CGCCGGTCGT GAAGGGACGG ACTGGGACCT CTGAAACTGG GluAlaMet AlaAlaSer ThrSerLeuPro AspProGly AspPheAsp> 200 180 GGAACGTGCC CCGGATCTGT GGGGTGTGTG GAGACCGAGC CACTGGCTTT 170 CCTTGCACGG GGCCTAGACA CCCCACACAC CTCTGGCTCG GTGACCGAAA ArgAsnValPro ArgIleCys GlyValCys GlyAspArgAla ThrGlyPhe> 230 CACTTCAATG CTATGACCTG TGAAGGCTGC AAAGGCTTCT TCAGGCGAAG GTGAAGTTAC GATACTGGAC ACTTCCGACG TTTCCGAAGA AGTCCGCTTC HisPheAsn AlaMetThrCys GluGlyCys LysGlyPhe PheArgArgSer> 280 CATGAAGCGG AAGGCACTAT TCACCTGCCC CTTCAACGGG GACTGCCGCA GTACTTCGCC TTCCGTGATA AGTGGACGGG GAAGTTGCCC CTGACGGCGT MetLysArg LysAlaLeu PheThrCysPro PheAsnGly AspCysArg> 330 TCACCAAGGA CAACCGACGC CACTGCCAGG CCTGCCGGCT CAAACGCTGT 320 AGTGGTTCCT GTTGGCTGCG GTGACGGTCC GGACGGCCGA GTTTGCGACA IleThrLysAsp AsnArgArg HisCysGln AlaCysArgLeu LysArgCys> 380 GTGGACATCG GCATGATGAA GGAGTTCATT CTGACAGATG AGGAAGTGCA CACCTGTAGC CGTACTACTT CCTCAAGTAA GACTGTCTAC TCCTTCACGT ValAspile GlyMetMetLys GluPhelle LeuThrAsp GluGluValGln>

	,			
410	420 * *	430 * *	· 440 * *	450 * *
GAGGAAGCGG	GAGATGATCC	TGAAGCGGAA	GGAGGAGGAG	GCCTTGAAGG
CTCCTTCGCC	CTCTACTAGG	ACTTCGCCTT	CCTCCTCCTC	CGGAACTTCC
ArgLysArg	GluMetIle I	euLysArgLys	GluGluGlu	AlaLeuLys>
		400	400	500
460 * *	470 * *	480 * *	490 * *	500 * *
ACAGTCTGCG	GCCCAAGCTG	TCTGAGGAGC	AGCAGCGCAT	CATTGCCATA
TGTCAGACGC	CGGGTTCGAC	AGACTCCTCG	TCGTCGCGTA	GTAACGGTAT
AspSerLeuArg	g ProLysLeu	SerGluGlu G	lnGlnArgIle	e IleAlaIle>
510	520 * *	530 * *	540 * *	550 * *
CTGCTGGACG	CCCACCATAA	GACCTACGAC	CCCACCTACT	CCGACTTCTG
GACGACCTGC	GGGTGGTATT	CTGGATGCTG	GGGTGGATGA	GGCTGAAGAC
LeuLeuAsp A	AlaHisHisLy	ThrTyrAsp	ProThrTyr	SerAspPheCys>
560	570	580	590	600
* *	* *	* *	* *	* *
CCAGTTCCGG	CCTCCAGTTC	GTGTGAATGA	TGGTGGAGGG	AGCCATCCTT
GGTCAAGGCC	GGAGGTCAAG	CACACTTACT	ACCACCTCCC	TCGGTAGGAA
GlnPheArg	ProProVal	ArgValAsnAsı	GlyGlyGly	SerHisPro>
610	620	630	640	650
* *	* *	* *	* *	* *
CCAGGCCCAA	CTCCAGACAC	ACTCCCAGCT	TCTCTGGGGA	CTCCTCCTCC
GGTCCGGGTT	GAGGTCTGTG	TGAGGGTCGA	AGAGACCCCI PheSerGlvAs	p SerSerSer>
Serargeroas	II SelAighis	THE TODGE		
660	670	680	690	700
* *	* *	* *	* *	* *
TCCTGCTCAG	ATCACTGTAT	CACCTCTTCA	GACATGATGG	ACTCGTCCAG
AGGACGAGTC	TAGTGACATA	GTGGAGAAGT	CTGTACTACC	TGAGCAGGTC
SerCysSer	AspHisCysIl	e ThrSerSer	AspMetMet	AspSerSerSer>
710	720			
* *		* *		* *
CTTCTCCAAI	CTGGATCTGA	. GTGAAGAAGA	TTCAGATGAC	CCTTCTGTGA
GAAGAGGTTA	GACCTAGACT	CACTTCTTCT	AAGTCTACTG	GGAAGACACT
PheSerAsn	LeuAspLeu	SerGluGluAs	p seraspast	PIOSELVAIP
760	770			
* 1		* *	* * *	
CCCTAGAGCT	GTCCCAGCT(TCCATGCTGC	CCCACCTGGC	TGACCTGGTC
GGGATCTCGA	A CAGGGTCGAC	, SerMetter	ProHistone	ACTGGACCAG La AspLeuVal>
ThrLeuGluLe	en pereture	1 Detwerpen	. TOUTSHEAD.	
810	=		840	850
	••	•		TACCAGGATT
AGTTACAGC	A TCCAAAAGG	L CHITGGCIII	CGATTCTAC	r ATGGTCCTAA
SerTurser	IleGlnLvsV	al IleGlyPhe	a AlaLysMet	IleProGlyPhe>
201131261			_	-

CAGAGACCTC ACTCTGAGG ACCAGATCT ACTGCTGAAG TCAAGTGCCA GTCTCTGGAG TGGAGACTCC TGGTCTAGCA TGACGACTTC AGTTCACGGT ArgAspLeu ThrSerglu AspGlnIleVal LeuLeuLys SerSerAla> 910 920 930 940 950 TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGACATG AACTCCAGTA GTACAACGCC AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPhethrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC	860	870	880	890	900
GTCTCTGGAG TGGAGACTCC TGGTCTAGCA TGACGACTTC AGTTCACGGT ArgAspLeu ThrSerGlu AspGlnIleVal LeuLeuLys SerSerAla> 910 920 930 940 950 TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGACATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPhethrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC					
910 920 930 940 950 TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGACATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACACA TACGCGCGTCA GTGACGTGAC					
TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGACATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC					
TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC	ArgAspLeu	ThrSerGlu A	AspGlnIleVal	l LeuLeuLys	SerSerAla>
TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC					
TTGAGGTCAT CATGTTGCGC TCCAATGAGT CCTTCACCAT GGACGACATG AACTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuarg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 1000 TCCTGGACCT GTGGCCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC	910				950
ARCTCCAGTA GTACAACGCG AGGTTACTCA GGAAGTGGTA CCTGCTGTAC IleGluValIle MetLeuArg SerAsnGlu SerPheThrMet AspAspMet> 960 970 980 990 1000 TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC					* *
Ser Ser					
TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC					
TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC	IleGluValIle	e MetLeuArg	SerAsnGlu S	SerPheThrme	c AspAspMet>
TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC	0.60	070	000		1000
TCCTGGACCT GTGGCAACCA AGACTACAAG TACCGCGTCA GTGACGTGAC		. * *	_		
AGGACCTGGA CACCGTTGGT TCTGATGTTC ATGGCGCAGT CACTGCACTG		GTGGCA A CCA			
SerTrpThr CysGlyAsnGln AspTyrLys TyrArgVal SerAspValThr> 1010 1020 1030 1040 1050 * * * * * * * * * * * * * * * * * * *					
CAAAGCCGGA CACAGCCTGG AGCTGATTGA GCCCCTCATC AAGTTCCAGG GTTTCGGCCT GTGTCGGACC TCGACTAACT CGGGGAGTAG TTCAAGGTCC LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle LysPheGln> 1060 1070 1080 1090 1100 ** * * * * * * * * * * * * * * * * *					
CAAAGCCGGA CACAGCCTGG AGCTGATTGA GCCCCTCATC AAGTTCCAGG GTTTCGGCCT GTGTCGGACC TCGACTAACT CGGGGAGTAG TTCAAGGTCC LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle LysPheGln> 1060 1070 1080 1090 1100 ** * * * * * * * * * * * * * * * * *	Set II pint	-ysgrynsiigr	r wahilina	Tythigvar	JCIMSDVGIIII >
CAAAGCCGGA CACAGCCTGG AGCTGATTGA GCCCCTCATC AAGTTCCAGG GTTTCGGCCT GTGTCGGACC TCGACTAACT CGGGGAGTAG TTCAAGGTCC LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle LysPheGln> 1060 1070 1080 1090 1100 ** * * * * * * * * * * * * * * * * *	1010	1020	1030	1040	1050
GTTTCGGCCT GTGTCGGACC TCGACTAACT CGGGGAGTAG TTCAAGGTCC LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle 1060 1070 1080 1090 1100 * * * * * * * * * * * * * * * * * *					
GTTTCGGCCT GTGTCGGACC TCGACTAACT CGGGGAGTAG TTCAAGGTCC LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle 1060 1070 1080 1090 1100 * * * * * * * * * * * * * * * * * *	CAAAGCCGGA	CACAGCCTGG	AGCTGATTGA	GCCCCTCATC	AAGTTCCAGG
LysAlaGly HisSerLeu GluLeuIleGlu ProLeuIle LysPheGln> 1060 1070 1080 1090 1100 * * * * * * * * * * * * * * * * * *					
TGGGACTGAA GAAGCTGAAC TTGCATGAGG ACCTGACTT CTTCGACTTG ACCTGACTT CTTCGACTTG ACCTGACTT CTTCGACTTG ACCTGACTT CTTCGACTTG CTCCTCGTACA GGACGAGTAC CTCTGTACA GGACGAGTAC CTCTCGTACA GGACGAGTAC 1110 1120 1130 1140 1150 * * * * * * * * * * * * * * * * * *					
TGGGACTGAA GAAGCTGAAC TTGCATGAGG AGGAGCATGT CCTGCTCATG ACCCTGACTT CTTCGACTTG AACGTACTCC TCCTCGTACA GGACGAGTAC ValGlyLeuLys LysLeuAsn LeuHisGlu GluGluHisVal LeuLeuMet> 1110 1120 1130 1140 1150 GCCATCTGCA TCGTCTCCCC AGATCGTCCT GGGGTGCAGG ACGCCGCGCT CGGTAGACGT AGCAGAGAGGGG TCTAGCAGGA CCCCCACGTCC TGCGGCGGGA AlaIleCys IleValSerPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 GCTGCCGCCA CCCGCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGGT GGGCGGGGC CCGTCGGTG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT					
TGGGACTGAA GAAGCTGAAC TTGCATGAGG AGGAGCATGT CCTGCTCATG ACCCTGACTT CTTCGACTTG AACGTACTCC TCCTCGTACA GGACGAGTAC ValGlyLeuLys LysLeuAsn LeuHisGlu GluGluHisVal LeuLeuMet> 1110 1120 1130 1140 1150 * * * * * * * * * * * * * * * * * * *	1060	1070	1080	1090	1100
ACCCTGACTT CTTCGACTTG AACGTACTCC TCCTCGTACA GGACGAGTAC ValGlyLeuLys LysLeuAsn LeuHisGlu GluGluHisVal LeuLeuMet> 1110 1120 1130 1140 1150 * * * * * * * * * * * * * * * * * * *					
ValGlyLeuLys LysLeuAsn LeuHisGlu GluGluHisVal LeuLeuMet> 1110 1120 1130 1140 1150 * * * * * * * * * * * * * * * * * * *					
1110 1120 1130 1140 1150 * * * * * * * * * * * * * * * * * * *					
GCCATCTGCA TCGTCTCCCC AGATCGTCCT GGGGTGCAGG ACGCCGCGCT CGGTAGACGT AGCAGAGGGG TCTAGCAGGA CCCCACGTCC TGCGGCGCGA AlalleCys IleValSerPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 GCTGCCGCCA CCCGCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * * * * * * *	ValGlyLeuLy	s LysLeuAsn	LeuHisGlu (GluGluHisVa	l LeuLeuMet>
GCCATCTGCA TCGTCTCCCC AGATCGTCCT GGGGTGCAGG ACGCCGCGCT CGGTAGACGT AGCAGAGGGG TCTAGCAGGA CCCCACGTCC TGCGGCGCGA AlalleCys IleValSerPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 GCTGCCGCCA CCCGCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * * * * * * *					
GCCATCTGCA TCGTCTCCCC AGATCGTCCT GGGGTGCAGG ACGCCGCGCT CGGTAGACGT AGCAGAGGGG TCTAGCAGGA CCCCACGTCC TGCGGCGCGA AlalleCys IleValserPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 * * * * * * * * * * * * * * * * * *					1150
CGGTAGACGT AGCAGAGGGG TCTAGCAGGA CCCCACGTCC TGCGGCGCGA AlaIleCys IleValSerPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 * * * * * * * * * * * * * * * * * *					* *
AlaileCys IleValSerPro AspArgPro GlyValGln AspAlaAlaLeu> 1160 1170 1180 1190 1200 * * * * * * * * * * * * * * * * * *					
1160 1170 1180 1190 1200 * * * * * * * * * * * * * * * * * *					
GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * *	AlalleCys	llevalserPr	o Aspargrio	GIYVAIGIN .	Asparaarabeu>
GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * *	1160	. 1170	1100	1190	1200
GATTGAGGCC ATCCAGGACC GCCTGTCCAA CACACTGCAG ACGTACATCC CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * *	1100				
CTAACTCCGG TAGGTCCTGG CGGACAGGTT GTGTGACGTC TGCATGTAGG IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * *	CAMMCACCC			САСАСТССАС	ACGTACATCC
IleGluAla IleGlnAsp ArgLeuSerAsn ThrLeuGln ThrTyrIle> 1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * GCTGCCGCCA CCCGCCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * * * * * * CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT	CTA ACTCCCC	TAGGTCCTGG	CGGACAGGTT	GTGTGACGTC	TGCATGTAGG
1210 1220 1230 1240 1250 * * * * * * * * * * * * * * * * * * GCTGCCGCCA CCCGCCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * * * * * * *	TIAGIUAla	TleGlnAsp	ArgLeuSerAs	n ThrLeuGln	ThrTyrIle>
GCTGCCGCCA CCCGCCCCG GGCAGCCACC TGCTCTATGC CAAGATGATC CGACGGCGGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * * * * * * *	1100101.10	22002.2.05	· · · · g • • · · · · · · · · · · · · ·		-
* * * * * * * * * * * * * * * * * * *	1210	1220	1230	1240	1250
CGACGCCGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT					* *
CGACGCCGT GGGCGGGGC CCGTCGGTGG ACGAGATACG GTTCTACTAG ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * * CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT	GCTGCCGCCA	CCCGCCCCC	GGCAGCCACC	TGCTCTATGC	CAAGATGATC
ArgCysArgHis ProProPro GlySerHis LeuLeuTyrAla LysMetIle> 1260 1270 1280 1290 1300 * * * * * * * * * * * CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT	CGACGGCGGT	GGGCGGGGG	CCGTCGGTGG	ACGAGATACG	GTTCTACTAG
1260 1270 1280 1290 1300 * * * * * * * * * * * * CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT	ArgCysArgHi	s ProProPro	GlySerHis	LeuLeuTyrAl	a LysMetIle>
* * * * * * * * * * * * * * * * * * *					
CAGAAGCTAG CCGACCTGCG CAGCCTCAAT GAGGAGCACT CCAAGCAGTA GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT					
GTCTTCGATC GGCTGGACGC GTCGGAGTTA CTCCTCGTGA GGTTCGTCAT					
	CAGAAGCTAG	CCGACCTGCC	CAGCCTCAAT	GAGGAGCACT	CCAAGCAGTA

12/20 1340 1330 CCGCTGCCTC TCCTTCCAGC CTGAGTGCAG CATGAAGCTA ACGCCCCTTG GGCGACGGAG AGGAAGGTCG GACTCACGTC GTACTTCGAT TGCGGGGAAC ArgCysLeu SerPheGln ProGluCysSer MetLysLeu ThrProLeu> 1380

1370 1360

TGCTCGAAGT GTTTGGCAAT GAGATCTCCT GA ACGAGCTTCA CAAACCGTTA CTCTAGAGGA CT ValLeuGluVal PheGlyAsn GluIleSer ***>

TRANSCRIPT 10 FIGURE 7 (Sequence Range: 1 to 1534) 30 20 MetGluTrpArg AsnLysLys> 90 80 AGGAGCGATT GGCTGTCGAT GGTGCTCAGA ACTGCTGGAG TGGAGGGGAT 70 TCCTCGCTAA CCGACAGCTA CCACGAGTCT TGACGACCTC ACCTCCCCTA ArgSerAsp TrpLeuSerMet ValLeuArg ThrAlaGly ValGluGlyMet> 130 GGAGGCAATG GCGGCCAGCA CTTCCCTGCC TGACCCTGGA GACTTTGACC CCTCCGTTAC CGCCGGTCGT GAAGGGACGG ACTGGGACCT CTGAAACTGG GluAlaMet AlaAlaSer ThrSerLeuPro AspProGly AspPheAsp> 200 180 GGAACGTGCC CCGGATCTGT GGGGTGTGTG GAGACCGAGC CACTGGCTTT 170 CCTTGCACGG GGCCTAGACA CCCCACACAC CTCTGGCTCG GTGACCGAAA ArgAsnValPro ArgIleCys GlyValCys GlyAspArgAla ThrGlyPhe> 230 220 CACTTCAATG CTATGACCTG TGAAGGCTGC AAAGGCTTCT TCAGGTGAGC GTGAAGTTAC GATACTGGAC ACTTCCGACG TTTCCGAAGA AGTCCACTCG HisPheAsn AlaMetThrCys GluGlyCys LysGlyPhe PheArg*** 280 CCCCCTCCCA GGCTCTCCCC AGTGGAAAGG GAGGGAGAAG AAGCAAGGTG GGGGAGGGT CCGAGAGGGG TCACCTTTCC CTCCCTCTTC TTCGTTCCAC 330 TTTCCATGAA GGGAGCCCTT GCATTTTCA CATCTCCTTC CTTACAATGT AAAGGTACTT CCCTCGGGAA CGTAAAAAGT GTAGAGGAAG GAATGTTACA 400 380 CCATGGAACA TGCGGCGCTC ACAGCCACAG GAGCAGGAGG GTCTTGGCGA 370 GGTACCTTGT ACGCCGCGAG TGTCGGTGTC CTCGTCCTCC CAGAACCGCT PCT/AU98/00817 WO 99/16872

	14	/20		440	450	
	420	4	30		* *	
AGCATGAAGC		*	* ~~~~~~	AACG GGGA	CTGCCG	
* *		ATTCACCT	GC CCCIIC	TTGC CCCT	GACGGC	
* * AGCATGAAGC C	GAAGGCTGA	TAAGTGGA	CG GGGAAG			
TCGTACTTCG (Chicorr			490	200	
		_	400		*	
460 * *	*	* *	* CCCCT	GCCGG CTC	AAACGCT	
* *	CACAACCGA	C GCCACTG	CCA GGCC1	CGGCC GAG	TTTGCGA	
* * CATCACCAAG GTAGTGGTTC	CTCTTGGCT	G CGGTGAC	GGT CCGO.		~50	
GTAGTGGTTC	CIGIL		_	540	550 * *	
			530	_	*	
210	*	* *	TOTA TTCT	GACAGA TG	AGGAAGTG	
510 * * GTGTGGACAT CACACCTGT	CGGCATGA	TG AAGGAG	AAGA	CTGTCT AC	TCC1.ICAC	
GTGTGGACAT	GCCGTACT	AC TTCCTC	AAGI 1210		600	
CACACCTGIA	, 00-		580	590	. *	
56	0 5	570		* *	76	
20	0 * *C GGGAGAT	* *	CCCC AAGO	GAGGAGG A	3GCC.T.T.Gvv	
~~ aanna	* * GGGAGAT	GAT CCTGA	MCCCC TTC	CTCCTCC T	CCGGAACII	
CAGAGGAAG	G CCCTCTA	CTA GGAC'I	TCGCC		650	1
GTC-TCC1 TC	,0		630		* *	•
6.	10	620	*	* *		A
*	* *	*	TGAGGA GCI	AGCAGCGC A	TA CTA A CGG	r
CACAGTO	10 * * TG CGGCCC AC GCCGGG	AAGC TGTC	ACTCCT CG	rcgrcgcg .	ragina.	
CCTCTCAC	AC GCCGGG	TTCG ACAG	ACTO	700	70	0
	660	670	* *	*	CTCCGACTT	rC 27
*	* *	AAG	CTACG AC	CCCACCIA	GAGGCTGA	A G
TACTGCT	GGA CGCCCI	ACCAI TETC	TGGATGC TO	3GGG,I.GGw.		
TACTGCT	GGA CGCCCI CCT GCGGG	TGGTA TTC	* * ACCTACG AC	3GGGTGGA1	7	50
ATGACGA	CC1 C		730	740	*	*
ATGACGA	710	720	730	740	*	*
ATGACGA	710	720	730	740	*	*
ATGACGA	710	720	730	740	*	*
ATGACGA	710	720	730 * * GTGTGAAT C	740	* GGAGCCAT CCTCGGTA	*
ATGACGA * TGCCAG ACGGTC	710 * TTCC GGCCT	720 * rccagt tcc aggtca agg	730 * ** GACACTTA C	740 * ATGGTGGAC TACCACCTC	GGAGCCAT CCTCGGTI	800 *CC
ATGACGA * TGCCAG ACGGTC	710 * TTCC GGCCT	720 · * rccagt tcc aggtca agg	730 * * * * * * * * * * * * * * * * * *	740 * ATGGTGGAG TACCACCTO 79	GGAGCCAT CCTCGGT	* CCC AGG 800 *
ATGACGA * TGCCAG ACGGTC	710 * TTCC GGCCT	720 · * rccagt tcc aggtca agg	730 * * * * * * * * * * * * * * * * * *	740 * ATGGTGGAG TACCACCTO 79	GGAGCCAT CCTCGGT	* CCC AGG 800 *
ATGACGA * TGCCAG ACGGTC	710 * TTCC GGCCT	720 · * rccagt tcc aggtca agg	730 * * * * * * * * * * * * * * * * * *	740 * ATGGTGGAG TACCACCTO 79	GGAGCCAT CCTCGGT	* CCC AGG 800 *
ATGACGA * TGCCAG ACGGTC	710 * TTCC GGCCT	720 · * rccagt tcc aggtca agg	730 * * * * * * * * CACACTTA 780 * * CACTCCCAG GTGAGGGTC	740 * ATGGTGGAG TACCACCT 79 * CTTCTCTGG GAAGAGAC	GGAGCCAT CCTCGGTA GGACTCCT GGACTCCT CCTGAGGA	* CCC AGG 800 * CCCT AGGA 850
ATGACGA * TGCCAG ACGGTC * TTCCA	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG	720 * TCCAGT TCC AGGTCA AGG 770 * * TCCAGAC AGAGGTCTG TCAGAC	730 * * * * * * CACACTTA 780 * * * CACTCCAG GTGAGGGTC	740 * * * * * * * * * * * * * * * * * *	GGAGCCAT CCTCGGTI CCTCGGTI GGACTCCT CCCTGAGGI	* CCC 800 * CCCT AGGA 850 *
ATGACGA * TGCCAG ACGGTC * TTCCA	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG	720 * TCCAGT TCC AGGTCA AGG 770 * * * TCCAGAC AGAGGTCTG TCC 820	730 * * * * * * CACACTTA 780 * * CACTCCAG GTGAGGGTC 830	740 * ATGGTGGAG TACCACCT 79 * CTTCTCTGG GAAGAGACC 8	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA	* CCC 800 * CCCT AGGA 850 *
ATGACGA * TGCCAG ACGGTC * TTCCA	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG	720 * TCCAGT TCC AGGTCA AGG 770 * * * TCCAGAC AGAGGTCTG TCC 820	730 * * * * * * CACACTTA 780 * * CACTCCAG GTGAGGGTC 830	740 * ATGGTGGAG TACCACCT 79 * CTTCTCTGG GAAGAGACC 8	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA	* CCC 800 * CCCT AGGA 850 *
ATGACGA * TGCCAG ACGGTC * TTCCA	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG	720 * TCCAGT TCC AGGTCA AGG 770 * * * TCCAGAC AGAGGTCTG TCC 820	730 * * * * * * CACACTTA 780 * * CACTCCAG GTGAGGGTC 830	740 * ATGGTGGAG TACCACCT 79 * CTTCTCTGG GAAGAGACC 8	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA	* CCC 800 * CCCT AGGA 850 *
* TGCCAG ACGGTC * TTCCA AAGGT	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG * * * * * * * * * * * * * * * * * *	730 * * ETGTGAAT G CACACTTA C 780 * * CACTCCCAG ETGAGGGTC 830 * ATCACCTCTT PAGTGGAGAA	740 * * *ATGGTGGAG TACCACCTC 79 * CTTCTCTGG GAAGAGACC 8 CAGACATG GTCTGTAC	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA A0 AT GGACTC TA CCTGAG	* CCC 800 * CCCT AGGA 850 * CGTCC GCAGG
* TGCCAG ACGGTC * TTCCA AAGGT	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG * * * * * * * * * * * * * * * * * *	730 * * * * * * 780 780 * * * * * * * * * * * * *	740 * ** ** ** ** * * * * * * * * * * * *	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA A0 AT GGACTC TA CCTGAC	* CCC AGG 800 * CCCT AGGA 850 * CGTCC GCAGG
* TGCCAG ACGGTC * TTCCA AAGGT	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG * * * * * * * * * * * * * * * * * *	730 * * * * * * 780 780 * * * * * * * * * * * * *	740 * ** ** ** ** * * * * * * * * * * * *	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA A0 AT GGACTC TA CCTGAC	* CCC AGG 800 * CCCT AGGA 850 * CGTCC GCAGG
* TGCCAG ACGGTC * TTCCA AAGGT	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG * * * * * * * * * * * * * * * * * *	730 * * * * * * 780 780 * * * * * * * * * * * * *	740 * ** ** ** ** * * * * * * * * * * * *	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA A0 AT GGACTC TA CCTGAC	* CCC AGG 800 * CCCT AGGA 850 * CGTCC GCAGG
* TGCCAG ACGGTC * TTCCA AAGGT	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG * * * * * * * * * * * * * * * * * *	730 * * * * * * 780 780 * * * * * * * * * * * * *	740 * ** ** ** ** * * * * * * * * * * * *	GGAGCCAT CCTCGGTA GGACTCCT CC CTGAGGA A0 AT GGACTC TA CCTGAC	* CCC AGG 800 * CCCT AGGA 850 * CGTCC GCAGG
ATGACGA * TGCCAG ACGGTC * TTCCA AAGGT CCTCC GGAG AGCC	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGE GACGAG TCT 860 * TTCTCCA AT AAGAGGT TCT	720 * TCCAGT TCC AGGTCA AGG * * *CCAGAC AGAGGTCTG TC * * *ATCACTGT AGAGTGACA TCAGTGACA TCAGACA TCAGACA TCAGACA TCAGACA TCAGACCTAGA	730 * * ETGTGAAT G CACACTTA C 780 * * CACTCCCAG ETGAGGGTC 830 * ATCACCTCTT TAGTGGAGAA GAGTGAAGA CTCACTTCT	740 * * ** ** ** ** * * * * * * * * * * *	GGAGCCAT CCTCGGTN CCCTCGGTN CCCTGAGGN A0 AT GGACTCCTACTACCTGAGGN AT GGACTCCTACCTGAGGN AT GGACTCCTACCTACCTACCTACCTACCTACCTACCTACCT	* CCC AGG 800 * CCT AGGA 850 * CGTCC CCAGGG 900 * TCTGT AGACA 950
ATGACGA * TGCCAG ACGGTC * TTCCA AAGGT CCTCC GGAG AGCC	710 * TTCC GGCCC AAGG CCGG 760 * GGCCC AACC CCGGG TTG: 810 * CTGCTC AGE GACGAG TCC 860 * TTCTCCA AT AAGAGGT TE	720 * TCCAGT TCC AGGTCA AGG 770 * * *CCAGAC AGAGGTCTG TCC 820 * * *ATCACTGT AGAGTGACA TCAGTGACA TCAGTGACA TCAGACTGT AGACCTAGA	730 * * CACACTTA C 780 * * CACTCCCAG CTGAGGGTC 830 * TCACCTCTT TAGTGGAGAA GAGTGAAGA CTCACTTCT 91	740 * ** * *ATGGTGGAG TACCACCT 79 * * * * * * * * * * * * * * * * *	GGAGCCAT CCTCGGTA CCTGAGGA CCTGAGGA AO AT GGACTCCT AT GGACTC CTA CCTGAG AT GCACTC AT CCTGAG AT C	* CCC AGG 800 CCCT AGGA 850 CCCC SCAGG 900 TCCCC AGGACA 950 * * * * * * * * * * * * * * * * * * *
ATGACGA * TGCCAG ACGGTC * TTCCA AAGGT CCTCC GGAG AGCC	710 * TTCC GGCCC AAGG CCGG 760 * GGCCC AACC CCGGG TTG: 810 * CTGCTC AGE GACGAG TCC 860 * TTCTCCA AT AAGAGGT TE	720 * TCCAGT TCC AGGTCA AGG 770 * * *CCAGAC AGAGGTCTG TCC 820 * * *ATCACTGT AGAGTGACA TCAGTGACA TCAGTGACA TCAGACTGT AGACCTAGA	730 * * CACACTTA C 780 * * CACTCCCAG CTGAGGGTC 830 * TCACCTCTT TAGTGGAGAA GAGTGAAGA CTCACTTCT 91	740 * ** * *ATGGTGGAG TACCACCT 79 * * * * * * * * * * * * * * * * *	GGAGCCAT CCTCGGTA CCTGAGGA CCTGAGGA AO AT GGACTCCT AT GGACTC CTA CCTGAG AT GCACTC AT CCTGAG AT C	* CCC AGG 800 CCCT AGGA 850 CCCC SCAGG 900 TCCCC AGGACA 950 * * * * * * * * * * * * * * * * * * *
ATGACGA * TGCCAG ACGGTC * TTCCA AAGGT CCTCC GGAG AGCC	710 * TTCC GGCCT AAGG CCGG 760 * GGCCC AACT CCGGG TTG: 810 * CTGCTC AGA GACGAG TCT	720 * TCCAGT TCC AGGTCA AGG 770 * * *CCAGAC AGAGGTCTG TCC 820 * * *ATCACTGT AGAGTGACA TCAGTGACA TCAGTGACA TCAGACTGT AGACCTAGA	730 * * CACACTTA C 780 * * CACTCCCAG CTGAGGGTC 830 * TCACCTCTT TAGTGGAGAA GAGTGAAGA CTCACTTCT 91	740 * ** * *ATGGTGGAG TACCACCT 79 * * * * * * * * * * * * * * * * *	GGAGCCAT CCTCGGTA CCTGAGGA CCTGAGGA AO AT GGACTCCT AT GGACTC CTA CCTGAG AT GCACTC AT CCTGAG AT C	* CCC AGG 800 CCCT AGGA 850 CCCC SCAGG 900 TCCCC AGGACA 950 * * * * * * * * * * * * * * * * * * *

960	970	980	. 990 * *	1000
* *				
			TTGCTAAGAT	
AGTCAATGTC	GTAGGTTTTC	CAGTAACCGA	AACGATTCTA	CTATGGTCCT
1010	1020	1030	1040	1050
1010	1020	* *	1040 * *	* *
			GTACTGCTGA	
			CATGACGACT	
AAGICICIGG	AGIGGAGACI	CCIGGICIMO	Cilidicolor	10110110100
1060	1070	1080	1090	1100
* *	* *	* *	. * *	* *
CATTGAGGTC			GTCCTTCACC	
			CAGGAAGTGG	
1110	1120	1130	1140	1150
* *	* *	* *	1140 * *	* *
			AGTACCGCGT	
			TCATGGCGCA	
1160	1170	1180	1190	1200
* *	* *	* *	1190 * *	* *
ACCAAAGCCG	GACACAGCCT	GGAGCTGATT	GAGCCCCTCA	TCAAGTTCCA
TGGTTTCGGC	CTGTGTCGGA	CCTCGACTAA	CTCGGGGAGT	AGTTCAAGGT
1210		1230	1240	1250
* *			* *	
GGTGGGACTG	AAGAAGCTGA	ACTTGCATGA	GGAGGAGCAT	GTCCTGCTCA
CCACCCTGAC	TTCTTCGACT	TGAACGTACT	CCTCCTCGTA	CAGGACGAGT
1260	1270	1280	1290	1300
* *			* *	
TGGCCATCTG	CATCGTCTCC	CCAGATCGTC	CTGGGGTGCA	CCTCCCCCCC
ACCGGTAGAC	GTAGCAGAGG	GGTCTAGCAG	GACCCCACGT	CCIGCGGCGC
4740	1220	1220	1340	1350
1310	1320	* * *	1340	* *
			AACACACTGC	
CTGATTGAGG	CCATCCAGGA	CCCCCIGICC	TTGTGTGACG	TCTGCATGTA
GACTAACICC	GGIAGGICCI	GGCGGMCAGG	11010101010	
1360	1370	1380	1390	
* *		1380	* *	* *
CCCCTCCCGC	CACCCGCCC	CGGGCAGCCA	CCTGCTCTAT	GCCAAGATGA
GGCGACGGCG	GTGGGCGGG	GCCCGTCGGT	GGACGAGATA	CGGTTCTACT
1410	1420	1430	1440	1450
* *	k * 1	. * *	* *	* *
TCCAGAAGCT	AGCCGACCT	G CGCAGCCTCA	ATGAGGAGCA	CTCCAAGCAG
AGGTCTTCG/	A TCGGCTGGA	C GCGTCGGAGI	TACTCCTCGT	GAGGTTCGTC
	-			
1460			1490	1500 * *
*	* *			,
TACCGCTGC	C TCTCCTTCC	A GCCTGAGTGC	AGCATGAAGC	TAACGCCCCT
ATGGCGACG	G AGAGGAAGG	r CGGACTCACC	TCGTACTTCG	ATTGCGGGGA

1520 *

TGTGCTCGAA GTGTTTGGCA ATGAGATCTC CTGA ACACGAGCTT CACAAACCGT TACTCTAGAG GACT

FIGURE 8	TRANSO	CRIPT 11		
10	20 *	30	40	50 *
			GGTGAGGCCA CCACTCCGGT	
60 *	70 *	80	90 *	100
			AACAGCGGCA TTGTCGCCGT	
110 *	120	130	140	150 *
			CCAATCCATC GGTTAGGTAG	
160 *	170 *	180	190	200
			GAGCCTGCCA CTCGGACGGT	
210	220	230	240	250 *
			GTGAGACCTC CACTCTGGAG	
260 *	270 *	280	290	300
			TCCTTCAGGG AGGAAGTCCC	
310	320	330	340	350 *
	GTGAAGGGAC	GGACTGGGAC	GAGACTTTGA CTCTGAAACT GlyAspPheAs	GGCCTTGCAC
360	370	380	, 390	400
	CACCCCACAC	ACCTCTGGCT	GCCACTGGCT CGGTGACCGA AlaThrGly	
410	420	430	440	450 *
ACGATACTGG	ACACTTCCGA	CGTTTCCGAA	CTTCAGGCGA GAAGTCCGCT e PheArgArg	TCGTACTTCG
460	470 *	480	490	500 *
CCTTCCGTGA	TAAGTGGACG	GGGAAGTTGC	GGGACTGCCG CCCTGACGGC GlyAspCysAr	CATCACCAAG GTAGTGGTTC g lleThrLys>

	_			
510	520	530	540	550
CTGTTGGCTĠ	GCCACTGCCA CGGTGACGGT	CCGGACGGCC	GAGTTTGCGA	CACACCTGTA
AspAsnArg A	ArgHisCysGlr	AlaCysArg	LeuLysArg C	ysValAspIle>
560 *	570 *	580 *	590 *	600 *
GCCGTACTAC	AAGGAGTTCA TTCCTCAAGT LysGluPhe l	AAGACTGTCT	ACTCCTTCAC	GTCTCCTTCG
610 *	620	630	640	650 *
CCCTCTACTA	CCTGAAGCGG GGACTTCGCC e LeuLysArg	TTCCTCCTCC	TCCGGAACTT	CCTGTCAGAC
660 *	670 *	680 *	690 +	700 *
GCCGGGTTCG	TGTCTGAGGA ACAGACTCCT LeuSerGluGlu	CGTCGTCGCG	TAGTAACGGT	TACTGCTGGA ATGACGACCT leLeuLeuAsp>
710	720	730	740	750
GCGGGTGGTA	AAGACCTACG TTCTGGATGC LysThrTyr	TGGGGTGGAT	GAGGCTGAAG	ACGGTCAAGG
760	770	780 *	790 *	800
CCGGAGGTCA	TCGTGTGAAT AGCACACTTA l ArgValAsn	CTACCACCTC	CCTCGGTAGG	TTCCAGGCCC AAGGTCCGGG SerArgPro>
810	820	830	840	850 *
TTGAGGTCTG	ACACTCCCAG TGTGAGGGTC HisThrProSe	GAAGAGACCC	CTGAGGAGGA	CCTCCTGCTC GGAGGACGAG SerSerCysSer>
860	870	880	890 *	900
TCTAGTGACA	T ATCACCTCTT A TAGTGGAGAA S IleThrSer	GTCTGTACTA	CCTGAGCAGG	TCGAAGAGGT
910	920	930	940	950 *
TAGACCTAGA	T GAGTGAAGAA A CTCACTTCTT	CTAAGTCTAC	TGGGAAGACA	CTGGGATCTC
AsnLeuAspLe	eu SerGluGlu	AspSerAsp	AspProSerVa	l ThrLeuGlu>
96	*	*	*	*
GACAGGGTC	C TCTCCATGCT G AGAGGTACGA LeuSerMetLe	CGGGGTGGAC	CGACTGGACC	TCAGTTACAG AGTCAATGTC ValSerTyrSer>

CATCCAAAAG GTCATTGGCT TTGCTAAGAT GATACCAGGA TTCAGAGACC GTAGGTTTTC CAGTAACCGA AACGATTCTA CTATGGTCCT AAGTCTCTGG IleGlnLys ValIleGly PheAlaLysMet IleProGly PheArgAsp>

TCACCTCTGA GGACCAGATC GTACTGCTGA AGTCAAGTGC CATTGAGGTC AGTGGAGACT CCTGGTCTAG CATGACGACT TCAGTTCACG GTAACTCCAG LeuThrSerGlu AspGlnIle ValLeuLeu LysSerSerAla IleGluVal>

1130

ATCATGTTGC GCTCCAATGA GTCCTTCACC ATGGACGACA TGTCCTGGAC TAGTACAACG CGAGGTTACT CAGGAAGTGG TACCTGCTGT ACAGGACCTG IleMetLeu ArgSerAsnGlu SerPheThr MetAspAsp MetSerTrpThr>

1180

CTGTGGCAAC CAAGACTACA AGTACCGCGT CAGTGACGTG ACCAAAGCCG GACACCGTTG GTTCTGATGT TCATGGCGCA GTCACTGCAC TGGTTTCGGC CysGlyAsn GlnAspTyr LysTyrArgVal SerAspVal ThrLysAla>

1230

GACACAGCCT GGAGCTGATT GAGCCCCTCA TCAAGTTCCA GGTGGGACTG CTGTGTCGGA CCTCGACTAA CTCGGGGAGT AGTTCAAGGT CCACCCTGAC GlyHisSerLeu GluLeuIle GluProLeu IleLysPheGln ValGlyLeu>

1280

AAGAAGCTGA ACTTGCATGA GGAGGAGCAT GTCCTGCTCA TGGCCATCTG TTCTTCGACT TGAACGTACT CCTCCTCGTA CAGGACGAGT ACCGGTAGAC LysLysLeu AsnLeuHisGlu GluGluHis ValLeuLeu MetAlaIleCys>

1330

CATCGTCTCC CCAGATCGTC CTGGGGTGCA GGACGCCGCG CTGATTGAGG GTAGCAGAGG GGTCTAGCAG GACCCCACGT CCTGCGGCGC GACTAACTCC IleValSer ProAspArg ProGlyValGln AspAlaAla LeuIleGlu>

1400 1380

CCATCCAGGA CCGCCTGTCC AACACACTGC AGACGTACAT CCGCTGCCGC GGTAGGTCCT GGCGGACAGG TTGTGTGACG TCTGCATGTA GGCGACGGCG AlaIleGlnAsp ArgLeuSer AsnThrLeu GlnThrTyrIle ArgCysArg>

CACCCGCCC CGGGCAGCCA CCTGCTCTAT GCCAAGATGA TCCAGAAGCT GTGGGCGGGG GCCCGTCGGT GGACGAGATA CGGTTCTACT AGGTCTTCGA HisProPro ProGlySerHis LeuLeuTyr AlaLysMet IleGlnLysLeu>

AGCCGACCTG CGCAGCCTCA ATGAGGAGCA CTCCAAGCAG TACCGCTGCC TCGGCTGGAC GCGTCGGAGT TACTCCTCGT GAGGTTCGTC ATGGCGACGG AlaAspLeu ArgSerLeu AsnGluGluHis SerLysGln TyrArgCys>

1510 1520 1530 1540 1550

TCTCCTTCCA GCCTGAGTGC AGCATGAAGC TAACGCCCCT TGTGCTCGAA AGAGGAAGGT CGGACTCACG TCGTACTTCG ATTGCGGGGA ACACGAGCTT LeuSerPheGln ProGluCys SerMetLys LeuThrProLeu ValLeuGlu>

1560 1570

GTGTTTGGCA ATGAGATCTC CTGA CACAAACCGT TACTCTAGAG GACT ValPheGly AsnGlulleSer ***>